

DOI:10.19651/j.cnki.emt.2107998

输送带损伤多尺度特征交叉融合检测方法*

屈鼎然¹ 乔铁柱^{1,2} 庞宇松³

(1. 太原理工大学 新型传感器与智能控制教育部重点实验室 太原 030024; 2. 太原理工大学 物理与光电工程学院 太原 030024; 3. 荷兰代尔夫特理工大学 机械海事和材料工程学院 代尔夫特 2628)

摘要: 基于图像或声音的单信号输送带纵向撕裂的检测方法往往受光线和噪声影响较大,为了克服这种局限性,提出了一种多尺度特征交叉融合检测方法。首先利用 Log-Mel 特征提取算法将一维声音信号转化为二维谱图;然后搭建双输入神经网络对图像与声谱图同时进行特征提取以及多尺度特征交叉融合;最后根据融合后特征通过损失函数判定损伤分类。经实验该方法的检测准确率、划痕和纵向撕裂的灵敏度分别可以达到 97.37%、96.53% 和 98.67%, 分别比现有的视听决策级融合检测方法提高了 7.04%、6.96% 和 6.4%。因此,该方法能够更好地满足输送带损伤检测的可靠性要求。

关键词: 视听;多尺度特征;交叉融合;输送带;损伤检测

中图分类号: TN911.7 文献标识码: A 国家标准学科分类代码: 510.40

Multi scale feature cross fusion detection method for conveyor belt damage

Qu Dingran¹ Qiao Tiezhu^{1,2} Pang Yusong³

(1. Key Laboratory of Advanced Transducers and Intelligent Control System, Ministry of Education, Taiyuan University of Technology, Taiyuan 030024, China; 2. College of Physics and Optoelectronics, Taiyuan University of Technology, Taiyuan 030024, China; 3. Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, CD Delft 2628, the Netherlands)

Abstract: The detection method of longitudinal tear of single signal conveyor belt based on image or sound is often affected by light and noise. In order to overcome this limitation, a damage detection method based on multi scale feature cross fusion is proposed. Firstly, Log-Mel feature extraction algorithm is used to upgrade one-dimensional sound signal to two-dimensional spectrum. Secondly, a dual input neural network is built to extract the features of the image and the sound spectrogram at the same time, and they are cross fused at different scales. Finally, according to the fused features, the damage classification is determined by the loss function. Through experiments, the detection accuracy, sensitivity of scratch and longitudinal tear of this method can reach 97.37%, 96.53% and 98.67% respectively, which is 7.04%, 6.96% and 6.4% higher than the existing audio-visual decision level fusion method. Therefore, this method can better meet the reliability requirements of conveyor belt damage detection.

Keywords: audio-visual; multiscale feature; signal fusion; conveyor belt; damage detection

0 引言

在矿用带式输送机的工作过程中,金属尖锐部件意外坠落等事故可能会使输送带发生纵向撕裂^[1-2],从而导致机器损坏、生产停滞,甚至人员伤亡。因此,准确检测输送带损伤是一项必不可少的任务^[3]。

与传统检测方法相比,如电磁检测法^[4]、超声波检测法等^[5-6],机器视觉检测技术具有更好的检测速度和精度^[7],无需接触被检测对象,并且维护成本更低^[8]。现有的基于机器视觉的输送带损伤检测方法包括:Chen 等^[9]提出了基

于 SVM 的监测技术来检测矿用输送带的损伤。Li 等^[10]使用优化的 SSR 算法检测输送带的损坏。Yu 等^[11]提出了一种基于中红外光和长红外光组合的双波段机器视觉方法来检测输送带损坏。Hou 等^[12]使用多光谱图像采集传感器将可见光分解为可见光、中红外光和远红外光,并将多光谱图像集成用于输送带纵向撕裂检测。Qiao 等^[13]提出了一种基于红外和可见光融合的集成双目视觉检测方法,用于检测矿用输送带的撕裂。Qu 等^[14]提出了一种基于自适应深度卷积网络的输送带损伤检测方法。

收稿日期:2021-10-01

* 基金项目: NSFC-山西煤基低碳联合基金(U1810121)、2020 年中央引导地方科技发展资金项目(YDZX2020140001796)资助

然而,由于成像原理,基于红外图像的方法在实际实验中仍然存在噪声大、效果差等缺陷,尤其是红外成像类设备成本极高。而由于输送带工作环境中光线很暗,能见度低,基于可见光视觉的方法依旧具有很大的局限性。为了克服这种局限性,Hou 等^[13]将声音信号引入到基于机器视觉的输送带纵向撕裂检测任务中,提出了一种输送带纵向撕裂视听融合检测方法。根据输送带的图像信号与声音信号分别进行检测,将两种检测结果在决策层融合后对输送带纵向撕裂与否进行最终诊断。然而,在决策级融合之前,该方法仍存在单信号受噪声影响较大的局限性,比如:图像受光线强弱影响较大,声音受噪声影响较大。因此,本文延用了 Hou 等^[13]的视听融合思想,结合深度学习算法,提出了一种输送带损伤多尺度特征交叉融合检测方法。

由于一维声音信号的特征具有少而难以提取的特点,所以本文首先采用 Log-Mel 特征提取算法将一维声音信号转化为二维图谱信号^[16],从而获得更加全面的声音特征。然后搭建双输入神经网络,将图像和声谱图的特征同时进行提取,并在不同尺度下进行交叉融合;最后将融合后的特征通过损失函数来确定损伤状况。多尺度特征交叉融合检测方法可以融合不同尺度的视听信号特征,用更复杂的特征级融合代替决策级融合,使检测模型具有更好的抗噪性与准确度。

本研究收集了输送带损伤视听信号数据集,训练了能够准确检测输送带纵向撕裂和划伤的神经网络模型。经对比实验验证,与视听融合检测方法相比,检测效果有极大地提高。

1 多尺度特征交叉融合检测模型

1.1 信号预处理

视觉信号采用被压缩成 416×416 大小的 RGB 图像。

声音信号首先需要被转换为声谱图,然后再被压缩为相同大小的二维图像。本研究中所使用的声谱图是具有 Mel 特征的二维谱图。为了提取 Mel 特征,本文首先通过 FFT 运算提取了声音的频谱信息,将声音的时域转换为频域,得到了以时间为横轴、频率为纵轴的复杂二维阵列;然后对先前得到的阵列求平方;接着进行 Mel 滤波,将音频转换到 Mel 频域;最后为了增强声音的低频表现,进行了 Log 运算,从而最终获得了声音的 Log-Mel 声谱图特征。输送带纵向撕裂声音样本的波形图如图 1(a)所示,声谱图如图 1(b)所示。

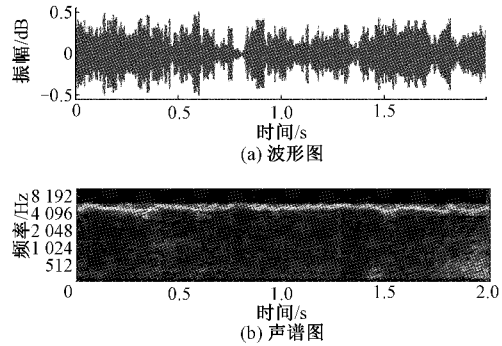


图 1 纵向撕裂声音样本的波形图与声谱图

1.2 神经网络结构

本文搭建了双输入端深度卷积网络,将输送带的图像和声谱图输入各自的主干网络进行特征提取。网络结构如图 2 所示。

图 2“CBL($x \times y \times m \times n // 2$)”中, x 是卷积核的深度, y 是卷积核的数量, $m \times n$ 是卷积核的大小,2 是步长,(步长默认为 1)。CBL 模块包含 1 个 2D 卷积层、1 个批量归一化层和 1 个 Leaky_ReLU 激活函数,激活函数如式(1)所示。

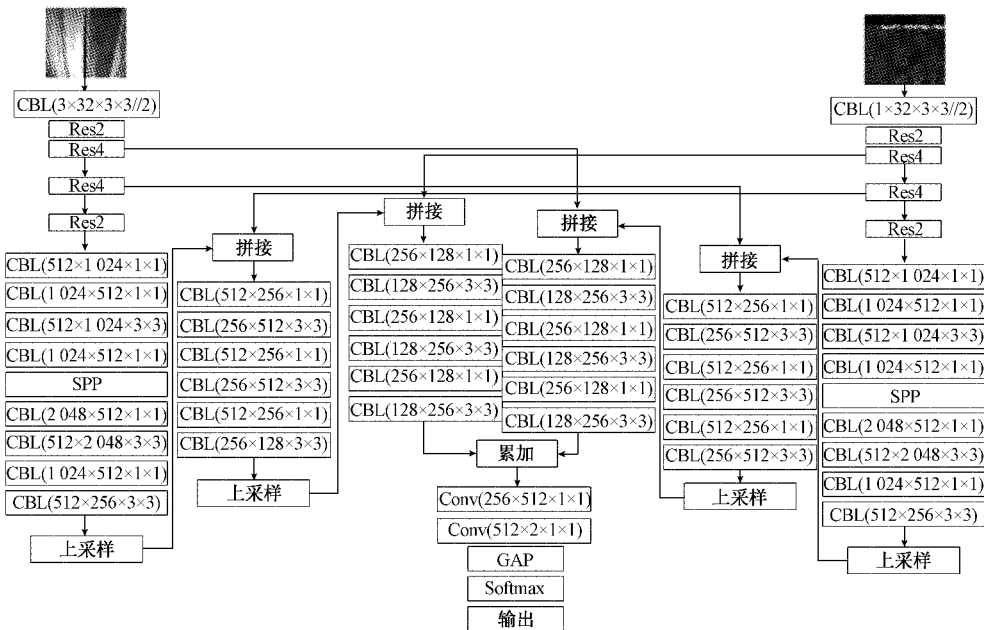


图 2 网络结构

$$y_i = \begin{cases} x_i, & x_i \geq 0 \\ \frac{x_i}{2}, & x_i < 0 \end{cases} \quad (1)$$

本文在网络主干中加入了残差模块 Res 以进行特征提取,其思想是将基础层的特征映射分为两部分,然后通过跨阶段的层次结构进行融合,优化网络中梯度信息重复的问题,不仅减少了计算量和内存开销,而且保证了检测精度。Res2 和 Res4 都包含了卷积核大小为 3×3,步长为 2 的卷积块,接下来 Res2 有 2 个残差结构,Res4 有 4 个。每个残差块的第 1 层卷积具有下采样的功能,因此每个 Res2 或 Res4 都会将特征图的比例缩小至 1/2。

本文还将 SPP 模块应用于网络模型^[17],使之可以提取出更加抽象的特征。SPP 模块可以有效地增加主干特征接收范围,并显著地分离出最重要的上下文特征。

1.3 多尺度特征交叉融合思想

为了融合输送带图像和声谱图的特征,本文提出了多尺度特征交叉融合方法。利用两种信号各自的特征提取网络提取两类图像的特征,然后将得到的不同尺度的特征图交叉进行拼接运算,最后累加。

这种融合方法对图像和声谱图中大小为 13×13 的特征图进行上采样,然后与对方大小为 26×26 的特征图进行拼接。之后采用 6 个 CBL 模块对特征图进行处理,以消除上一次上采样的混叠效应。之后再次进行上采样,然后将其与对方大小为 52×52 的特征图进行拼接,并通过 6 个 CBL 模块进行处理。最后对双方进行累加运算,进行最终的特征信息融合。

该方法将不同尺度特征图进行交叉融合,因此可以联系两种特征图的上下文特征,通过小尺寸特征图获得更加丰富的特征语义信息,从大比例尺特征图中获得更准确的目标位置信息。通过多次的交叉迭代融合,获得视听信号融合紧密、抽象性好、语义完善的特征。多尺度特征交叉融合示意图如图 3 所示。

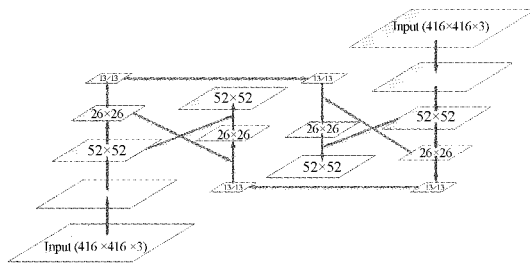


图 3 多尺度特征交叉融合思想

1.4 GAP+Softmax 输出结构

本文采用全局平均池化 GAP 和 Softmax 损失函数作为网络结构的输出模型^[18-19],减少了计算量和神经网络模型的规模,并且避免了过拟合。本文对视听信号特征进行融合,经过最后一个卷积层获得了两个 52×52 大小的特征图,分别表示划痕与纵向撕裂。然后,对每个特征图进

行 GAP 运算,即分别计算每个特征图特征点的平均值,作为分类的置信值。

GAP+Softmax 输出结构如图 4 所示,G1 和 G2 是划伤和纵向撕裂两个检测类别的置信值,S1 和 S2 是最终分类概率。

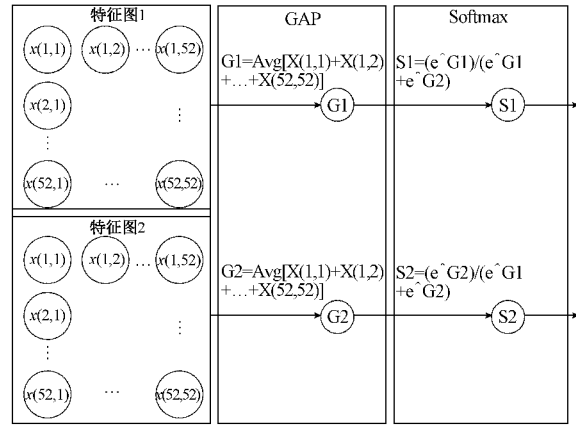


图 4 GAP+Softmax 输出结构

2 实验过程以及结果分析

2.1 实验平台搭建

1) 硬件平台搭建

为了验证多尺度特征交叉融合检测方法的可靠性,在本文中搭建了矿用输送带试验台。输送带长、宽、厚分别为 11 m、0.8 m、15 mm。矿用输送带实验平台如图 5 所示。

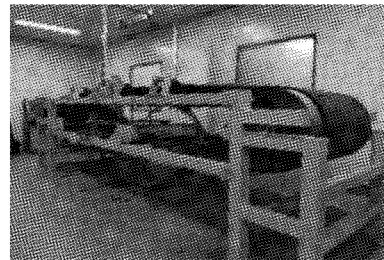


图 5 矿用输送带实验平台

本文设计了一种视听传感器,用于同时采集输送带的视听信号。该传感器主要包括 1 个拾音器和 1 个分辨率为 1 920×1 080,帧率为 24 fps 的可见光摄像机。

由于输送带的纵向撕裂通常发生在落煤点附近,因此本文选择在落煤点稍靠前处安装视听传感器。另外,由于输送带的下垂弧度,纵向撕裂更容易在下部表面被发现,并且输送带下部空间大,且不会被煤炭覆盖。因此本文将传感器安装在传送带的下部。

为了采集输送带损伤信号,本文在传感器附近安装了用于划伤输送带的金属螺栓。实验平台的硬件结构如图 6 所示。

2) 软件平台搭建

本实验采用的软件环境版本为:python3.6、tensorflow-gpu2.2.0rc2、keras2.4.3、numpy1.19.1、opencv3.3.1。

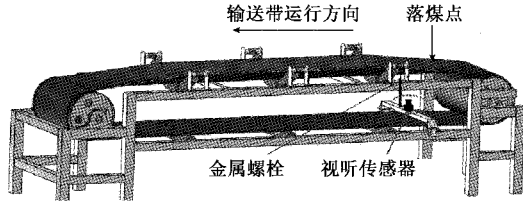


图 6 安装位置示意

2.2 数据集

在本实验中使用金属螺栓对输送带下表面进行摩擦,模拟其被尖锐物体划伤撕裂的情况。然后,采集了输送带出现表面划伤以及穿透性纵向撕裂时的视觉图像和音频数据。其中两类样本各包含 4 320 幅 RGB 图像和 180 s 音频文件。本实验以 1 s 为单位对音频文件进行分割,并将分割后的音频模块转换成 Mel 谱图。另外,由于每一秒音频对应 24 帧图像文件,本试验选择第 13 帧图像作为对应样本。因此,最终获得了划伤与纵向撕裂视听样本各 180 对,它们被当作实验训练集。图 7 所示为划伤的数据集样本,图 8 所示为纵向撕裂的数据集样本。

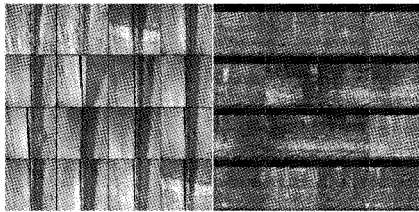


图 7 划伤的数据集样本

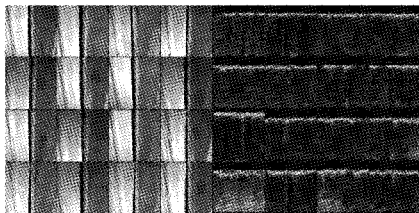


图 8 纵向撕裂的数据集样本

2.3 实验过程

1) 训练

为了获得输送带损伤检测模型,本实验将训练集输入神经网络进行训练。设置 epoch 为 1 000, batch 为 32。最终训练时间为大约 9 h。

2) 测试

为了更接近实际工况,本实验将带式输送机的运行速度设定为 3.15 m/s,并且模拟了暗光以及噪音环境。调整了金属螺栓的位置,采集了正常、划伤、纵向撕裂 3 种类型的视听样本,每种类型 600 组。损失函数输出概率阈值设置为 0.7,取概率最高的分类作为检测结果输出。如果无超过阈值概率的分类,则判定输送带状态正常。为了得到更准确的平均性能指标数据,在相同的实验条件下,连续进行了 5 次测试实验。

2.4 实验结果分析

为了与 Hou 等^[15]的方法形成对比,本文选取了一样的性能参数标准,即准确度和两种损伤类型的灵敏度。准确度是正确检测的样本总数(表 1)与实验中所有样本总数的比值。灵敏度是一种类型(表 1)中正确检测到的样本数与该类型(表 1)的实际样本总数之比。表 2 所示为这种方法的性能分析。

表 1 多尺度特征交叉融合方法结果

实验	正常	划伤	纵向撕裂
1	583	581	590
2	584	576	596
3	579	578	593
4	577	576	589
5	584	585	592

表 2 多尺度特征交叉融合方法性能分析 %

实验	准确度	划伤灵敏度	纵向撕裂灵敏度
1	97.44	96.83	98.33
2	97.56	96.00	99.33
3	97.22	96.33	98.83
4	96.78	96.00	98.17
5	97.83	97.50	98.67
平均	97.37	96.53	98.67

为了验证检测模型的可靠性,本文采用相同的实验数据集,在相同的暗光噪声环境下对 Hou 等^[15]提出的输送带纵向撕裂视听检测方法进行了测试。实验结果以及性能分析如表 3 和 4 所示。两种方法的最终性能比较结果如图 9 所示。

表 3 视听融合检测方法结果

实验	正常	划伤	纵向撕裂
1	535	531	551
2	505	527	547
3	548	544	553
4	559	548	560
5	528	537	557

表 4 视听融合检测方法性能分析 %

实验	准确度	划伤灵敏度	纵向撕裂灵敏度
1	89.83	88.50	91.83
2	87.72	87.83	91.17
3	91.39	90.67	92.17
4	92.61	91.33	93.33
5	90.11	89.50	92.83
平均	90.33	89.57	92.27

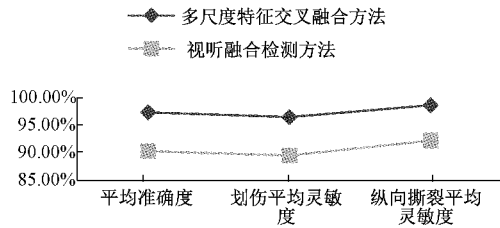


图9 平均性能比较

3 结 论

本文采用视听传感器采集矿用输送带损伤的二维可见光图像和一维音频信息,接着通过 Log-Mel 特征提取算法将声音转换为声谱图,然后构建了双输入神经网络,用于特征提取和多尺度特征交叉融合,最后通过 GAP 和 Softmax 损失函数获得最终的判定结果。这种方法可以更好地融合并利用视觉与声音信号的特征,克服了暗光、噪声条件下现有视听检测方法的局限性,拥有更高的检测精度。

本文搭建了矿用输送带损伤检测实验平台,采集了接近现场工况的实验数据集,训练了多尺度特征交叉融合检测模型。最后,在噪声实验环境下与 Hou 等^[15]的视听检测方法进行对比,多尺度特征交叉融合检测方法的平均准确率为 97.37%,比视听融合检测方法提高了 7.04%;划伤和纵向撕裂的检测灵敏度分别为 96.53%、98.67%,比视听融合检测方法分别提高 6.96% 和 6.4%。因此,多尺度特征交叉融合检测方法在暗光与噪音环境下依旧可以满足工业现场输送带损伤检测的可靠性要求。

在今后的研究中,本研究希望从如下几个方面进行改进,以完善输送带损伤多尺度特征交叉融合检测方法:1)收集更多复杂情况下的数据集,比如因托辊卡死造成的磨损情况、磨损时出现烟雾从而降低能见度的环境情况等,完善数据集使其更接近实际生产状况;2)通过实验找到更加合适的图像与声音的特征权重比,完善多尺度特征交叉融合模型;3)调整神经网络参数,继续提升精度或检测速度。

参考文献

[1] HAKAMI F, PRAMANIK A, RIDGWAY N, et al. Developments of rubber material wear in conveyor belt system [J]. Tribology International, 2017, 111: 148-158.

[2] YANG Y, HOU C, QIAO T, et al. Longitudinal tear early-warning method for conveyor belt based on infrared vision[J]. Measurement, 2019, 147: 106817.

[3] ANDREJIOVA M, GRINCOVA A, MARASOVA D. Measurement and simulation of impact wear damage to industrial conveyor belts [J]. Wear, 2016, 368: 400-407.

[4] 郭太平, 裘进浩, 程军, 等. 高频电磁涡流检测系统及实验研究[J]. 国外电子测量技术, 2015, 34(11):4-9.

[5] 李智, 闫玉波, 钟炜, 等. 塑封半导体器件特殊封装缺陷的声学扫描检测[J]. 国外电子测量技术, 2017, 36(10): 110-114.

[6] 闫蓓, 刘孟轲, 马宏林. 超声检测系统设计及回波信号特征分析 [J]. 电子测量技术, 2019, 42(2): 123-129.

[7] 郝勇, 耿佩, 温钦华, 等. 滚动轴承保持架缺陷的图像处理及模式识别方法研究[J]. 仪器仪表学报, 2019, 40(9): 162-169.

[8] FACCO P, SANTOMASO A C, BAROLO M. Artificial vision system for particle size characterization from bulk materials [J]. Chemical Engineering Science, 2017, 164: 246-257.

[9] CHEN G, ZHANG X, WANG Z J, et al. Robust support vector data description for outlier detection with noise or uncertain data [J]. Knowledge-Based Systems, 2015, 90: 129-137.

[10] LI J, MIAO CH Y. The conveyor belt longitudinal tear on-line detection based on improved SSR algorithm[J]. Optik-International Journal for Light & Electron Optics, 2016, 127(19): 8002-8010.

[11] YU B, QIAO T, ZHANG H, et al. Dual band infrared detection method based on mid-infrared and long infrared vision for conveyor belts longitudinal tear [J]. Measurement, 2018, 120: 140-149.

[12] HOU C, QIAO T, HAI Z, et al. Multispectral visual detection method for conveyor belt longitudinal tear[J]. Measurement, 2019, 143: 246-257.

[13] QIAO T, CHEN L, PANG Y, et al. Integrative binocular vision detection method based on infrared and visible light fusion for conveyor belts longitudinal tear[J]. Measurement, 2017, 110: 192-201.

[14] QU D, QIAO T, PANG Y, et al. Research on ADCN method for damage detection of mining conveyor belt[J]. IEEE Sensors Journal, 2021, 21(6): 8662-8669.

[15] HOU C, QIAO T, QIAO M, et al. Research on audio-visual detection method for conveyor belt longitudinal tear [J]. IEEE Access, 2019, 7: 120202-120213.

[16] MENG H, YAN T, YUAN F, et al. Speech emotion recognition from 3D Log-Mel spectrograms with deep learning network [J]. IEEE Access, 2019, 7: 125868-125881.

[17] HUANG Z, WANG J, FU X, et al. DC-SPP-YOLO:

- Dense connection and spatial pyramid pooling based YOLO for object detection[J]. Information Sciences, 2020, 522: 241-258.
- [18] HSIAO T, CHANG Y, CHOU H, et al. Filter-based deep-compression with global average pooling for convolutional networks [J]. Journal of Systems Architecture, 2019, 95: 9-18.
- [19] SUN Y, WEN G. Ensemble softmax regression model for speech emotion recognition[J]. Multimedia Tools & Applications, 2017, 76: 8305-8328.

作者简介

屈鼎然, 硕士研究生, 主要研究方向为神经网络和煤矿安全检测。

E-mail: goding234@163.com

乔铁柱, 教授, 主要研究方向为机器视觉、新型传感器和智能控制系统。

E-mail: qiaotiezhu@tyut.edu.cn

庞宇松, 教授, 主要研究方向为大型物料搬运系统和物流过程的智能控制。

E-mail: pystyut@163.com